DOCUMENT RESUME

ED 048 799 24 FL 002 154

AUTHOR	Weiss, Louis
TITLE	Auditory Discrimination and Identification in Foreign Language Learning.
INSTITUTION	Stanford Univ., Calif. Stanford Center for Research and Development in Teaching.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
REPORT NO	P-5-0252-0402; TR-17
BUREAU NO	BR-5-0252
PUB DATE	Feb 71
CONTRACT	OEC-6-10-078
NOTE	52p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Audition (Physiology), *Auditory Discrimination, Auditory Tests, Aural Stimuli, Conditioned Response, Discrimination Learning, *Educational Experiments, Identification Tests, Instructional Program Divisions, Language Instruction, Language Laboratories, *Language Research, Modern Languages, Morphophonemics, Phonemics, *Pronunciation, *Second Language Learning, Teaching Methods

ABSTRACT

The main purpose of this study was to investigate the validity of the assumption that auditory discrimination and pronunciation in a foreign language are closely related. If the assumption were to be well-founded, then it might be possible for foreign language students with high auditory discrimination ability to work alone in the language laboratory with appropriately structured materials and achieve near-native pronunciation. The design of the study involved developing and testing an instrument to measure auditory discrimination ability, which was followed by an investigation of the differential effects of two types of pronunciation training upon students with varying degrees of auditory discrimination ability as measured on that instrument. Some 200 subjects at the sixth-grade level participated in the experiment. Results indicate fundamental differences between mimicry and reproduction in second language learning. Consideration of the several perceptual tasks involved in responding to the stimuli presented in this study underscores the importance of distinguishing between the processes of discrimination and identification. Extensive use of statistical tables and charts illustrates experimental test results. (Author/RL)

ED048799

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

STANFORD CENTER
FOR RESEARCH AND DEVELOPMENT
IN TEACHING

Technical Report No. 17

AUDITORY DISCRIMINATION AND IDENTIFICATION IN FOREIGN LANGUAGE LEARNING

Louis Weiss

School of Education
Stanford University
Stanford, California

February 1971

FL 002 154

Published by the Stanford Center for Research and Development in Teaching, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the Office of Education. (Contract No. OE-6-10-078, Project No. 5-0252-0402.)

Introductory Statement

The Center is concerned with the shortcomings of teaching in American schools: The ineffectiveness of many American teachers in promoting achievement of higher cognitive objectives, in engaging their students in the tasks of school learning, and, especially, in serving the needs of students from low-income areas. Of equal concern is the inadequacy of American schools as environments fostering the teachers' own motivations, skills, and professionalism.

The Center employs the resources of the behavioral sciences--theoretical and methodological--in seeking and applying knowledge basic to achievement of its objectives. Analysis of the Center's problem area has resulted in three programs: Heuristic Teaching, Teaching Students from Low-Income Areas, and the Environment for Teaching. Drawing primarily upon psychology and sociology, and also upon economics, political science, and anthropology, the Center has formulated integrated programs of research, development, demonstration, and dissemination in these three areas. In the Heuristic Teaching area, the strategy is to develop a model teacher training system integrating components that dependably enhance teaching skill. In the program on Teaching Students from Low-Income Areas, the strategy is to develop materials and procedures for engaging and motivating such students and their teachers. In the program on Environment for Teaching, the strategy is to develop patterns of school organization and teacher evaluation that will help teachers function more professionally, at higher levels of morale and commitment.

Technical Report No. 17 attempts to help clarify a situation prevalent in language research and teaching which confounds the ability to discriminate (to detect differences between two sounds) and the ability to categorize (to identify each of two sounds as members of a class). The study also deals with the difference between mimicry (imitation of a model) and reproduction (recall removed in time from a model, which demonstrates storing and learning of sounds). The teaching of language presents special problems to teachers of disadvantaged ethnic groups. This study was conducted as part of the Teacher Training project of the Center's program on Teaching Students from Low-Income Areas.

Table of Contents

	Page
Introductory Statement	iii
List of Tables	vii
List of Figures	ix
Abstract	xi
Background	1
The Research Problem	3
Procedure	7
Design of the Study	7
Auditory Discrimination Test	11
Echo Response Test	14
Results and Discussion	18
Auditory Discrimination Test	19
Echo Response and Recall Tests	23
Data Analysis	25
Summary and Conclusions	35
References	40
Appendix A. Analysis of Variance of Scores on Lorge-Thorndike IQ Test, Auditory Discrimination Test and Echo Response #1	42
Appendix B. Auditory Discrimination Test	43
Appendix C. Answer Sheet for Auditory Discrimination Test	45

List of Tables

Table		Page
1	Auditory Discrimination Test: Summary Statistics of Total Test and Subtests	19
2	Summary Statistics of Echo Response #6	23
3	Summary Statistics of Recall Test	24
4	Summary Statistics of Analyses of Regression of ER #6 on Three Types of AD Test Scores: Both Treatment Groups	26
5	Summary Statistics of Analysis of Regression of ER #6 on Uncorrected AD Test Scores: Sex Within Treatment Groups	31
6	Three-way Analysis of Variance of Scores on Echo Response #6	32
7	Mean Scores on Echo Response #6 for All Groups	33
8	Three-way Analysis of Variance of Scores on Recall Test .	33
9	Mean Scores on Recall Test for All Groups	34

List of Figures

Figure		Page
1	Regression and Scatter Plot of Echo Response Test Scores on AD Test (Uncorrected) Scores: Tape Recorder Treatment and Teacher Treatment	28
2	Regression and Scatter Plot of Echo Response Test Scores on AD Test (Uncorrected) Scores: Tape Recorder Treatment--Male and Female	29
3	Regression and Scatter Plot of Echo Response Test Scores on AD Test (Uncorrected) Scores: Teacher Treatment--Male and Female	30

Abstract

The main purpose of this study was to investigate the validity of the assumption that auditory discrimination and pronunciation in a foreign language are closely related. If this assumption is well-founded, then it might be possible for foreign language students with high auditory discrimination ability to work alone in the language laboratory with appropriately structured materials and achieve criterion (near-native) pronunciation.

The design of the study involved developing and testing an instrument to measure auditory discrimination ability, which was followed by an investigation of the differential effects of two types of pronunciation training upon students with varying degrees of auditory discrimination ability as measured on that instrument. The two types of pronunciation training, designed to parallel language laboratory-based and teacher-based instruction, involved the oral presentation of 12 French vocabulary items to 200 subjects at the sixth-grade level. In the first training condition, stimuli were presented directly by the investigator or a trained assistant; in the second condition, stimuli were presented by the tape-recorded voice of the investigator or the assistant. In each condition, each item, accompanied by an appropriate visual illustration, was presented six times. An echo response was required of the subject after each presentation. Subjects' responses were recorded on a separate tape recorder and rated by a specially trained native speaker of French. The sixth Echo Response for each item comprised the dependent variable. After the fourth, eighth, and twelfth items, the illustrations, unaccompanied by any oral stimulus, were shown in the same order in which they had been originally presented and subjects were asked to recall the French name (Recall Test).

The hypothesis was that at different levels of auditory discrimination two treatments in pronunciation are differentially effective. To test it, scores on the sixth Echo Response were regressed on scores on the Auditory Discrimination (AD) Test. This regression analysis revealed no differential effects of the two treatments upon accuracy of pronunciation as measured by the instruments constructed for this study. The hypothesis was, therefore, not supported by the data.

Regression analyses and analyses of variance of scores for male and female subjects in each of the treatment groups did indicate an interaction of sex and auditory discrimination ability, as measured by the AD Test. The interaction contributed significantly to the variance of the scores on the sixth Echo Response. Analyses of scores on the Recall Test showed a small but consistent superiority of male and female subjects with high AD Test scores over their counterparts with low AD Test scores, and a superiority of female subjects over male subjects regardless of treatment group or AD Test scores.

The results indicate two theoretical distinctions of major importance to research on language learning. Comparison of the results on the Echo Response Test with those on the Recall Test reflects the crucial difference between mimicry and reproduction in foreign language learning. Consideration of the several perceptual tasks involved in responding to the stimuli presented in this study underscores the importance of distinguishing between the processes of discrimination and identification. These interrelationships have important implications with respect to research on the assumptions underlying the effective use of the language laboratory. If research in this field is to be productive and meaningful, it must be based on a better understanding of these distinctions.

AUDITORY DISCRIMINATION AND IDENTIFICATION IN FOREIGN LANGUAGE LEARNING

Louis Weiss[1]
Palo Alto Unified School District

The foreign language teacher today faces a dilemma. Never before in the history of American education has his profession enjoyed so much prestige and so much moral and financial support from the public, and yet never before has the profession been so uncertain of the theoretical and methodological foundations of language learning.

Background

To a great extent, this uncertainty has arisen because the surge in interest in foreign language study was accompanied by a quasi-official endorsement of the so-called "audio-lingual method" before certain of the assumptions on which it was based had been thoroughly investigated. The audio-lingual method emphasizes the listening and speaking aspects of language learning by concentrating on the memorization of dialogues meant to simulate real-life situations and by pattern practice involving the substitution of grammatically similar structural elements.

In its haste to implement the National Defense Education Act of 1958, the foreign language profession put an almost indelible stamp of approval on an approach to foreign language learning that had been used with some success in the Army Specialized Training Program during the Second World War and later, with modifications, in some elementary school classes. Since 1958, NDEA Foreign Language Institutes, with the cooperation of the Modern Language Association and many leading universities, have addressed themselves to the task of training and retraining secondary school foreign language teachers in this method. Observation and experience indicate that the implementation of the audio-lingual method has been somewhat less than universally successful.

---

[1]Formerly Research Assistant, Stanford Center for Research and Development in Teaching; now English teacher, Lewis Terman Junior High School.

In addition to dialogue memorization and structure drills, the audio-lingual method calls for an extended prereading period during which the student is completely dependent on the sound-system of the foreign language and on his ability to perceive, discriminate, retain, and reproduce utterances based on this unfamiliar sound-system. While the NDEA was training teachers in the use of the "New Key" and subsidizing the purchase of textbooks based on this method, authorities in the fields of linguistics and language teaching were raising serious questions about the validity of the assumptions underlying the use of dialogues, pattern practice, and extended prereading periods. Although studies such as the Colorado Experiment by Scherer and Wertheimer (1964) found positive results in favor of the audio-lingual method, many articles in such journals as International Review of Applied Linguistics, Language Learning, The French Review, and The Modern Language Journal, as well as the influential book by psychologist Wilga M. Rivers (1964), expressed serious doubts about the method. It is not surprising that in the midst of all this The Modern Language Journal published an article entitled "The Bewildered Modern Language Teacher" (Zeldner, 1963). That this bewilderment was not isolated or fleeting is evident from the fact that, in the last few years, three of the major language textbook companies have revised and modified the dialogue and drill components in their texts.

Of even greater concern to the profession and to administrators, since it involves large expenditures of public funds, is the tenuous status of the language laboratory. The language laboratory was considered by the proponents of the "New Key" to be a necessary and important adjunct to the classroom as support for the teacher whose pronunciation of the foreign language was below the acceptable level of proficiency, as a means of providing practice outside of regular class time and, especially, as a means of providing much-needed individualized instruction. So important were language laboratories considered to be that the NDEA provided matching funds to school districts for their purchase and installation. Between 1958 and 1966, the number of language laboratories in secondary schools in the United States increased from a few dozen to almost 10,000 (Hutchinson, 1966). Unfortunately, the laboratories were installed before

proper instructional materials could be developed and before teachers could be trained in the use of the electronic equipment. The many problems arising from these conditions are discussed by Hutchinson and are dramatized by two studies: the so-called Pennsylvania Project (Smith & Berger, 1968; Smith & Baranyi, 1968) and the Keating Report (Keating, 1963).

The rationale behind both studies was to investigate the effectiveness of the laboratory, not under ideal conditions, but under conditions actually existing in the field. The cumulative effect of these and similar studies was to add little to existing knowledge of the potential effectiveness of the language laboratory, and to lose support for it among teachers, administrators, and the general public.

The purpose of the present study was to investigate some of the assumptions underlying the use of the language laboratory and the audio-lingual method. The study was specifically concerned with assumptions regarding auditory discrimination and its relation to pronunciation in a foreign language.

The Research Problem

The assumption that in order to acquire accurate pronunciation, the learner must be able to distinguish between phonemes of the foreign language, as well as between those of the foreign language and his native language, is essential to the rationale of the "listening, speaking, reading, writing" sequence of learning set forth in the "New Key." Nelson Brooks (1964), one of the leading proponents of this view, states that the first immediate legitimate objective of second-language learning "is training in listening comprehension. Although language sounds originate in the voice box of the throat and are modified into recognizable speech movements in the mouth, it is the ear that dominates the learning and use of speech sound. Ear training must come first." This assumption is so widely held that auditory discrimination tests are frequently substituted for pronunciation tests (Lado, 1961), and in many cases, auditory discrimination training is used as a first step in pronunciation training.

This is also the rationale for the use of the language laboratory. In the laboratory, the student is supposed to be able to compare his pronunciation with that of the native speaker whose voice is on the tape, make the necessary successive approximations to the native speaker's pronunciation, and in time reach near-native fluency in the foreign language. This promise, however, has seldom been fulfilled. As Pimsleur points out,

> It is evident that students are rather poor judges of their own pronunciation. They are prone to think their pronunciation is good enough when actually it is not acceptable. They are unable to note which features are relevant and which are not. . . . If the language laboratory is to realize its full potentiality as a teaching machine . . . [the student] must be trained to correct his own speech. To do so, he must know which features are relevant and note how correct pronunciation sounds as distinct from incorrect pronunciation (1963, p. 199).

On the strength of these observations, Pimsleur and his associates conducted a series of experiments on auditory discrimination (Pimsleur, Mace, & Keislar, 1961). Seeking a refinement of the notion of discrimination, the experimenters provisionally adopted various meanings of the term "discrimination," prepared appropriate materials for each, and tried them out in 11 different pilot studies. The hypothesis under investigation in these studies was that discrimination training coupled with language laboratory practice would bring about better pronunciation than could laboratory practice alone. They found that, when the training involved French nasal vowel phonemes /õ/, /ã/, and /ẽ/, it was effective in producing good pronunciation of these phonemes. Training in discrimination between English diphthongized and French undiphthongized final vowel /o/, however, made no difference in the performances of the trained and untrained students; that is, the students who did not receive discrimination training performed as well as those who did receive the training. It should be noted that the first discrimination concerns distinctions that are phonemic in French but not in English, while the second concerns a French phoneme and the English mispronunciation of that phoneme. In both languages, the distinction between diphthongized and undiphthongized /o/ is nonphonemic, i.e., a phonetic distinction.

Pimsleur, in his article discussing this study (1963), speculates as follows:

> Evidently, these two pronunciation problems differ in some way so that discrimination training was effective for one and not the other. An examination of this difference should help clarify the role of discrimination. The following formulation is proposed for the reader's consideration: Learning the pronunciation of a foreign language means learning to make DIFFERENTIATED RESPONSES to DISCRIMINATIVE STIMULI (p. 201).

Pimsleur concludes that:

> discrimination training was effective for the French nasal vowels because the problem was mainly one of 'labeling,' that is, learning what phoneme classes exist in French and what words fit into each class. . . . On the other hand, discrimination training was ineffective with the French versus American o because the problem was not one of discrimination but of differentiation. Since this is a motor skill, it was improved by practice in making the necessary articulatory adjustments (p. 202).

In pointing out that discrimination training for the vowel differences was apparently irrelevant to improved pronunciation, Pimsleur draws a distinction between an auditory and a motor focus for pronunciation training. It is obvious that the two focuses are quite different and that they require quite different training procedures. The problem of their order of presentation, their relative importance, and their interaction is an extremely complex one.

This concept of phonetic differentiation and phonemic discrimination has important implications for the instructional use of the language laboratory and for the priorities established by the audio-lingual method. In the language laboratory, the student almost invariably spends his time listening to strings of sounds and practicing the imitation of these sounds until his pronunciation is "near-native." In many foreign language classes, students are not allowed to progress until they have mastered the technique (in French, for example) of producing an undiphthongized vowel or an unaspirated initial unvoiced plosive (i.e., /p/, /t/, /k/). The mastery of these phonetic details is not crucial to communication, since they are not phonemic and do not, thus, affect meaning. Such lack of

mastery is, rather, manifested in a "foreign accent"--a problem far less severe than inability to understand or be understood in a language. It is clear that both of these skills are relevant to language learning, and that both are closely related to the concept of "identity."

The concept of identity in the psychology of perception has been defined by Hebb (1949). He states that a figure (or, in this case, a sound) is perceived as having identity when it is recognized immediately as similar to some figures (or sounds) and dissimilar to others--that is, "when it falls into certain categories and not into others." The object (or sound) that does not have identity is recalled with great difficulty or not at all, and is not recognized or named easily. This identity, Hebb asserts, "depends on a considerable degree of experience; it is not innately given." Presumably, the categories that the beginning foreign language student has at his disposal comprise the phonemic system of his native language, but not that of the target language.

Harlan Lane (1966) points out that the acquisition of language involves mapping bits of phonic substance into distinct and discrete categories. It is Lane's belief that the less phonic substance the foreign language sound shares with a native sound, the more difficult is its identification. If a foreign language sound falls within the range of phonic substance of a native language sound (or shares some of its phonic substance), it is easily identified as that native sound.

Catford (1965) calls such identification of foreign sounds "phonological translation." It can occur receptively, he points out, and

> it occurs productively, though inadvertently, in the imperfect pronunciation of someone speaking a foreign language. We usually call this latter "phonological (or phonetic interference" and think of it as transference of native phonology into the foreign language being spoken (p. 61).

For students who are able to identify and discriminate foreign language sounds as such, the language laboratory can be an important and valuable adjunct to classroom learning. For students with little auditory discrimination ability, practice in the language laboratory could very well be frustrating, irritating, and dysfunctional. The problem to which

this study addressed itself, then, was to identify the level of auditory discrimination ability at which pronunciation practice in the language laboratory begins to be effective.

Procedure

This study is related to a preliminary investigation conducted by the experimenter in collaboration with Robert L. Politzer at the Stanford Center for Research and Development in Teaching. That study (Politzer & Weiss, 1969) examined the relationship of auditory discrimination, pronunciation, and recall to maturation. Testing students ranging in age from six to 14 years, the investigators found that achievement in the three language tasks increased significantly as a function of age. Since there is a danger in developmental studies of generalizing about cause-effect relationships in the presence of unknown and uncontrolled variables, no conclusions were drawn concerning the relationship of auditory discrimination and pronunciation. The fact that students improve consistently in pronunciation and in auditory discrimination does not necessarily mean that the two are related.

Design of the Study

Critical examination of the design of the previous study suggested that a more rigorous means of investigating the relationship of these language learning tasks would involve holding students' age constant and manipulating one or more of the critical variables. Since the question of the potential effectiveness of the language laboratory is central to the design of the present study, a simulated language laboratory condition was used in the manipulation of the pronunciation variable.

The design follows rather closely the Posttest-only Control Group Design (Number 6) as described by Campbell and Stanley (1963):

$$\begin{array}{llll} R & X_1 & O_1 \\ R & X_2 & O_2 \end{array}$$

where the Rs represent randomization, the Xs represent the exposure of two groups to two experimental pronunciation-practice treatments, and the Os

represent the scores achieved on the pronunciation test by the two groups. Scores of the two treatment groups on the dependent variable (pronunciation) and on the independent variable (auditory discrimination) are used to test the main research hypothesis: At different levels of auditory discrimination, two treatments in pronunciation training are differentially effective.

Scores on the independent variable (auditory discrimination) were based on a 29-item test designed to measure students' ability to discriminate between minimal pairs of French vowel sounds. The stimuli were presented to students by tape recorder and students marked their responses on an answer sheet.

Scores on the dependent variable (pronunciation) were based on students' echo responses to 12 French vocabulary items after a short period of pronunciation training. The two kinds of pronunciation training were designed to parallel language laboratory-based and teacher-based instruction. In the first training condition, which simulated teacher-based instruction (Teacher Treatment), stimuli were presented directly by the investigator or his assistant vis-à-vis the student. In the second condition, which simulated language laboratory-based instruction, the same stimuli were presented by the tape-recorded voice of the investigator or his assistant. In each condition, each item was presented six times, accompanied by an appropriate visual illustration of the item. After each presentation, the students were expected to respond with as exact an echo of the stimulus as possible. Stimuli and students' echo responses were later rated by a specially trained native speaker of French. Criterion for each response was the exact echo of the previous stimulus. Thus, the first five stimulus-response pairs were considered as pronunciation training while the sixth echo response was the measure of pronunciation ability.

After the fourth, eighth, and twelfth items, students were shown the visual illustrations in the same order in which they had been presented originally and were asked to recall the French name (Recall Test).

Description of the sample. The 200 students used in the study were selected from 11 sixth-grade classes in five elementary schools in a Bay Area school district. The decision to use sixth-grade students was based on the results of the preliminary investigation which indicated that the sixth grade seemed to be the point at which auditory discrimination (or the ability to take this type of auditory discrimination test) is developed enough to provide a measure that could be used as an independent variable. In addition, the necessity for using students with little or no previous exposure to foreign language study restricted the choice to the elementary school level since almost all students in that school district begin daily foreign language study in the seventh grade.

Since the schools were assigned by the district, the question arose whether there might be initial differences among the populations of the five schools that could influence or bias the results of the dependent variable. With this in mind, scores on the Auditory Discrimination Test, the first Echo Response, and the Lorge-Thorndike IQ were examined by analysis of variance, testing the differences among the 11 classes in the five schools, the differences between the two treatment groups, and the interaction of treatment and class. These analyses yielded no significant F-ratios and showed no indication of interaction of class and treatment groups. (These analyses of variance are summarized in Appendix A.)

The 200 students in the sample were randomly selected from among all the sixth-grade classes in the five schools. This selection was done by stratified random sampling on the basis of school and sex; that is, within each school, an equal number of boys and girls was selected from a table of random numbers. Six alternates, three boys and three girls, were selected from among the remaining students in case an accident in recording eliminated one or more of the original 200 students. In fact, three of the original students, two boys and one girl, had to be dropped because their speaking tests had been inadvertently erased or not recorded. Replacements were made from the emergency pool by random selection. A few students with speech and hearing difficulties (as identified by school records) were excluded from the random selection, as were five students who had indicated previous exposure to French, and who had veri-

fied this by answering questions in French put to them by the experimenter.

As in most studies of this kind, it is of great interest to consider the sex variable, since it is usually expected that girls will perform better than boys. Because of unequal distribution of boys and girls in the various classes, the assignment of students to sex-treatment groups yielded slightly unequal Ns, as follows:

	Male	Female
Treatment A (Tape Recorder)	52	48
Treatment B (Teacher)	48	52

Description of the instruments. Both the Auditory Discrimination Test (hereafter referred to as the AD Test) and the Echo Response Test (hereafter referred to as the ER Test) are adaptations of the instruments devised by Politzer with the assistance of the experimenter for the preliminary investigation in 1967-68. Although the ER Test remained essentially the same, it was necessary to change the AD Test to fit the purposes of the present study.

In the earlier study, students in the first, third, fifth, seventh, and ninth grades were presented with 48 items featuring 20 monosyllabic and disyllabic pairs of identical French words and 28 minimal pairs of French monosyllabic and disyllabic words (including eight pairs of French and English words resembling each other closely in pronunciation). The students' task consisted of indicating whether the members of each pair were the same or different. This task proved to be extremely easy; students in the fifth, seventh, and ninth grades achieved better than 90 percent correct responses. Thus, although the test was valuable for the purposes of the developmental study, it was not considered sensitive enough for the present study.

The only standardized test that seemed to approach the requirements of this study was the Sound Discrimination subtest of the Pimsleur Language Aptitude Battery (Pimsleur, 1966). This battery, however, was

considered too difficult for sixth-grade students. It was, therefore, necessary to write a new test of auditory discrimination appropriate to the sixth-grade level.

Auditory Discrimination Test

The AD Test was administered in the first week of March 1969, and the ER Test was administered two weeks later. An item analysis of the AD Test indicated that the test as it stood could not be used as an independent variable in this study. It was obvious from the proportion of correct responses that more than half the items were too easy (significances for these items were well above .80) and from the interscale correlations that the test lacked internal consistency. The Cronbach Alpha of .48 obtained from the analysis was a further indication of the need for revision and readministration. Consequently, several revisions were made, each revision being pretested with sixth-grade students in surrounding school districts. The final revision was completed early in May 1969 and readministered to the classes participating in the study. Since none of the items in the original test appeared in the final version, and since the format of the two versions was substantially altered, it was possible to tell the students that they were taking a different test.

The final version of the AD Test consisted of 29 items recorded on audiotape by the experimenter (see Appendix B). The recording was made on high quality tape-recording equipment at the Stanford University Speech Research Laboratory with the assistance of the director and the technical staff. The fidelity of the sound met the standards established by the laboratory. Each test item contained two contrasting French sounds embedded in so-called minimal pairs, that is, two strings of sounds which are identical except for the contrasting vowel phonemes. The items were presented in the form of ABX triads. The first two members of the triad, A and B, differed only in respect to the critical vowel phonemes and were separated by a pause of one second. The third member, X, was a replica of either A or B and was separated from B by a pause of two seconds. The student's task was to indicate on an answer sheet whether X sounded more

like A or B or whether he was not sure. Students were given four seconds after each item for marking their responses. Timing was controlled by tape equipment.

Before the recording session, the member of the pair AB that was to be used as the X member of the triad was selected at random and the items were then arranged by means of stratified random sampling in the following manner.

The first three items consisted of disyllabic strings in which the contrasting vowel sounds were placed in the first syllable. These items were ordered at random. The following seven items consisted of three-syllable strings with the contrasting vowel sounds embedded in the second syllable. These items were ordered at random and a pause was provided after the third set. The next nine items contained strings of four syllables with the contrasting vowel sounds in the second syllable. These items were ordered at random and a pause was provided after the fourth set. The last ten items were electronic reproductions of ten three-syllable and four-syllable items selected from the first part of the test. These items were also randomly ordered and a pause was provided after the fifth set. Three types of vowel contrasts were used:

1. Rounded vs. unrounded high front vowels (e.g., /y/ vs. /i/). In this contrast, both vowel sounds operate phonemically in French and are differentiated by the marked feature of rounding. This feature does not, however, operate phonemically in English, and American students are likely to perceive the rounded vowel sound as unrounded.

2. Nasal vs. non-nasal vowels (e.g., /æ̃/ vs. /æ/). In this contrast, as in the one above, both vowel sounds operate phonemically in French and are differentiated by the marked feature of nasality. Again, this latter feature does not operate phonemically for English vowels, and students are likely to perceive the nasal vowel as oral or as a nasalized vowel followed by "n."

3. Nasal vowels vs. other nasal vowels (e.g., /ã/ vs. /õ/). In this contrast, both vowel sounds operate phonemically in French, share the marked feature of nasality and differ by two features: position of

the tongue and lip-rounding. Two major factors contribute to the difficulty of this contrast. Non-nasal back vowels such as /a/ and /o/ are difficult even for speakers of English to distinguish and in many dialects tend to be produced and perceived in free variation. When, as in the case of the contrast presented here, attention is focused on the strongly marked feature of nasality, this vowel discrimination becomes even more difficult to make.

Since these three contrasts seem to be measuring different degrees of complexity in auditory discrimination, it seems reasonable to regard each type of contrast as a separate subtest of the entire auditory discrimination test. A summary of the types of vowel contrasts and the number of syllables in the various items is presented below.

Vowel Contrast	Number of Syllables in Each String 2	3	4	Total Contrast Type
rounded/unrounded	0	3	4	7
nasal/non-nasal	2	4	7	13
nasal vowel/ other nasal vowel	1	4	4	9

The entire tape script of the Auditory Discrimination Test, including the directions, is presented in Appendix B.

The AD Test was administered free field with a portable tape recorder operated by the experimenter to each of the 11 classes in their respective classes and in the presence of the teacher. There were rarely more than 25 students in a class, and care was taken to see that no one was seated too far from the tape recorder. Specially prepared answer sheets (see Appendix C) were distributed to the students, and the experimenter gave the directions for the test-taking procedure. Two practice items were used as examples and when there was little doubt that the students knew what was required of them, the classroom teacher stepped outside the door to prevent interruptions by messengers, tardy students, and so on. The experimenter then started the tape recorder and the test began.

Echo Response Test

The ER Test consisted of 12 French vocabulary items, each item comprised of a short noun and its definite article. In all but one case (le bateau), each item was two syllables in length, thus putting very little constraint on the students' memory. Since the experimental design required that the items be accompanied by a picture, it was necessary to restrict the choice of vocabulary to concrete nouns which could be unambiguously pictured. The pictures were drawn by a commercial artist in india ink on 8 by 10 Bristol board.

The items were selected because they contain the vowel phonemes that are considered most likely to cause pronunciation problems for American students. Five of the items contain monophthongal vowels which speakers of English tend to pronounce as diphthongs, namely: le sou, le lit, la clef, le seau, and le bateau. The last item (le bateau) also contains the French a, which English speakers often mispronounce as a schwa. Four of the items contain front rounded vowels which do not exist phonemically in English, namely: le mur, la feuille, le feu, and le coeur. The three remaining items contain examples of French nasal vowels, namely: le pont, l'agent, and le train. These items are particularly difficult since English speakers have a tendency to pronounce nasal vowels as non-nasal. Even when these vowels are pronounced as nasals, they are subject to interference from English and are often mispronounced as vowel plus "n."

In the test itself, the items were ordered from a table of random numbers and appeared in the following sequence:

1. le sou	7. la clef
2. le bateau	8. le feu
3. le lit	9. le seau
4. le mur	10. le train
5. la feuille	11. le pont
6. l'agent	12. le coeur

Each subject was shown a picture representing the item and was given a model of the pronunciation. The model consisted of six repetitions of the items as presented either by a tape recording or by the experimenter,

depending upon the treatment assigned to the subject. The items were separated by three-second pauses. The task of the subject was to repeat the pronunciation after each utterance, during the three-second pause. The pronunciation models and the subject's echo responses were recorded on a second tape recorder. For both treatments, the subject was asked to look at the pictures as he listened and repeated. However, after the third repetition, the picture was placed face down, since experience in the pilot study suggested that the pictures themselves distracted the subject from the main task of listening to the models and repeating. Three repetitions with the picture cue were considered sufficient to establish a bond between the French sounds and their referents. The sixth repetition was the criterion for the ER Test.

Since the Auditory Discrimination Test had been recorded by the experimenter, there was a possibility that similarity of speaking voice and style might bias the results on a test in which subjects are asked to mimic the pronunciation and intonation of a speaker. To avoid this possible bias, as well as any others that could easily arise from the use of a single experimenter, an assistant was used to administer the ER Test to approximately half of the subjects. The assistant was a female teacher with considerable experience in teaching French at the junior high school level and with a near-native command of the language. Before the actual administration of the test, she was trained by the experimenter and was given the opportunity to practice the testing technique with 15 fifth-grade pupils in an elementary school not involved in the study. Each experimenter made his own tape recording of the stimulus words in exactly the same manner as they were to be presented live. That is, the repetitions were not re-recorded electronically. Due to the inevitable variation that accompanies all speech, the six live repetitions of each item could be expected to be slightly different. Although this difference is almost imperceptible, it is nevertheless present. For purposes of experimental rigor, it was considered advisable to match the recorded utterances with the live ones. However, since the criterion for correct student pronunciation was similarity to the preceding model rather than to an arbitrary "acceptable" French pronunciation, the slight variations presented no problems.

It was necessary, however, to consider the problem of feedback from the experimenter. It is natural for a teacher to reward students for correct responses by either verbal or nonverbal reinforcement. In the language laboratory, the absence of the teacher deprives the student of this reinforcement. To provide such reinforcement in the experiment would be to add a very important variable not accounted for in the experimental design. On the other hand, to withhold any reinforcement from the experimenter would be to run the risk of losing rapport and discouraging the subject from trying his best. A compromise was reached by having the experimenter receive each repetition by the subject with a slight nod of his head, an action designed to be neutral, while at the same time being friendly and encouraging.

For the convenience of the participating schools, and to cause as little disruption to school routine as possible, only one experimenter was present at the administration of the ER Test. The principal experimenter tested three classes in one school and two classes in another, and the assistant experimenter tested the students in the remaining three schools. Male and female students as well as the number of subjects were divided equally among the two investigators.

In each of the schools, a room (usually the nurse's office or the teachers' workroom) was provided for the testing. Two tape recorders were set up on a table separating the experimenter from the subject. A microphone was placed in such a way that the subject could speak directly into it from his seat. Each classroom teacher was given a roster of the students being tested and was instructed to send the subjects into the testing room one at a time. The experimenter's copy of the roster contained the code numbers (1 through 200) for each student and a special code designating the treatment group to which the student had been assigned ("A" for Tape Recorder Treatment and "B" for Teacher Treatment). Before the student was admitted to the testing room, the experimenter recorded the code number but not the treatment designation on the student tape. This was done to eliminate rater bias, since it was not possible for the rater to hear the difference between the live model and the tape recorded model. (The rater was not aware of the two-treatment dimension of the study.)

After the student was comfortably seated, he was asked to record his name on the student tape. The experimenter then played the tape back to insure that the recorder was working properly and explained the testing procedure to the student. He then held up the first picture and began the appropriate modeling treatment. After the fourth, eighth, and twelfth words, the experimenter stopped the modeling and said: "Now let's see how many of these words you can remember." The pictures were then shown to the student in the order in which they had been presented in the Echo Response Test, and the student was allowed five seconds to recall the French word and record it on the tape. The entire testing procedure took between eight and ten minutes and all 200 students and six alternates were tested within a period of twelve school days.

The responses on the ER Test were judged by a native speaker of French (the wife of a foreign graduate student at Stanford University) who had served as a rater in the preliminary study by Politzer and Weiss. In that study, ratings had also been obtained from a second judge (who was also a native speaker of French) and the correlation of the independent ratings of these two judges was $r = .85$, an indication that the responses could be reliably rated. For purposes of the present experiment, the judge was retrained by the experimenter. Both the experimenter and the judge rated the performances of the 15 students who had been recorded during the training sessions for the assistant experimenter and reached a level of 85% agreement. During the rating of the actual tests, spot checks were made regularly by the experimenter to insure that this level of agreement was sustained.

Scores for each student were obtained from the first, third, and sixth response to each of the 12 words. The criterion for correctness was an acceptable duplication by the student of the relevant vowel sound in the preceding stimulus and each response was rated as either correct or incorrect. In cases where the vowel sound was correctly duplicated but a consonant was mispronounced, the response was rated as "correct." The sixth Echo Response was used as the main dependent variable.

Scores were also obtained for each student on his performance on the Recall Test. Criterion for correctness in this case was an utterance that

was clearly recognizable by the rater as the French name for the referent. Thus, even if the relevant vowel sound in a word was mispronounced, the word was still considered as correctly recalled.

All scores were punched on IBM cards and verified by professional key-punch operators. The data were processed by a computer.

Summary of testing procedure. The study was restricted to one age level, that of public school students in the sixth grade. The relationship of auditory discrimination and pronunciation was investigated by analyzing the interaction effects of auditory discrimination ability and two pronunciation-teaching treatments. Auditory discrimination ability was measured by a test developed for the experiment. In this test, subjects were presented with stimuli via tape recorder and were asked to mark their responses on an answer sheet. The two treatments took the form of an Echo Response Test embedded in what was ostensibly a French vocabulary learning test. This test was given to subjects on an individual basis. A subject was shown pictures of a series of objects and was told he was to learn the name of each object by listening to a model and repeating. The French name for each object was modeled six times, each utterance followed by an echo response from the subject. Treatment A presented the subject with an auditory model only, provided by tape recorder, thus requiring him to rely entirely on his auditory discrimination ability for pronunciation clues. Treatment B presented the subject with an auditory-visual model, provided by the experimenter, thus relieving him of complete reliance on his auditory discrimination ability. The subjects' responses were recorded on audiotape and rated by an independent judge. The sixth response on each item was used as the main dependent variable, i.e., the measure of pronunciation ability. A secondary dependent variable, ancillary to the main study but nevertheless interesting and important, was the student's ability to recall the French vocabulary used in the Echo Response Test.

Results and Discussion

Analysis of the data revealed no differential effects of auditory discrimination ability as measured by the AD Test constructed for this

study. The hypothesis was, therefore, not supported by the data. Further analysis revealed no differential effects of the two pronunciation training treatments as measured by the Echo Response Test constructed for this study. A detailed discussion of these negative results will serve to highlight some persistent problems involved in research on auditory discrimination.

Auditory Discrimination Test

Results of the Auditory Discrimination (AD) Test were obtained by means of a computer program designed to analyze either continuous or dichotomous data. The program computes reliability coefficients (Cronbach's Alpha), means, standard deviations, skewness, and kurtosis for the total test as well as for subtests. For purposes of item analysis of the AD Test, three subtests were designated: Subtest 1, comprised of seven items, contains the contrasts between rounded and unrounded vowels; Subtest 2, comprised of 13 items, contains the contrasts between nasal and non-nasal vowels; Subtest 3, comprised of nine items, contains the contrasts between nasal vowels and other nasal vowels. In this analysis, the data were treated as dichotomous, that is, a response was either correct or incorrect, with selection of the "Not Sure" option considered an incorrect response. Summary statistics of the total test and the three subtests are presented in Table 1.

TABLE 1

Auditory Discrimination Test: Summary Statistics of Total Test and Subtests

Tests	Alpha	Mean	sd	Skewness	Kurtosis
Total Test	.61	17.40	3.86	-0.39	.36
Subtest 1	.57	5.38	1.52	-1.01	.59
Subtest 2	.55	7.28	2.46	.00	-0.13
Subtest 3	.35	4.74	1.70	-0.05	-0.69

It is evident from the low alphas in Table 1 that scores on the total test as well as on the three subtests are not sufficiently reliable to be used without a great deal of caution. An item analysis indicated a lack of internal consistency that may be attributable to a lack of consistency in the contrasts presented within each subtest. That is, the difficulty of the task may vary from item to item. This possibility is perhaps best considered in terms of the concept of distinctive features (Jakobson, Fant, & Halle, 1952). Briefly, for purposes of discussion here, certain features, such as rounding, voicing, aspiration, nasality, and so forth, can be considered distinctive or nondistinctive in a given language. In English, for example, the sound represented by the orthographic -p- in spin is different from the initial sound in pin in that the former is unaspirated and the latter is aspirated. Speakers of English, however, cannot use the aspirated vs. the unaspirated /p/ to distinguish between two words, as can speakers of Hindi, for example. Thus, the feature of aspiration is considered nondistinctive in English. The feature of voicing, on the other hand, can be considered distinctive in English, since the presence or absence of voicing may be the only cue which serves to distinguish one word from another. That is to say, the b- of bin is different from the p- of pin in that the speaker of English perceives the /b/ and /p/ as two different sounds, causing a difference in meaning. Many linguists believe that speakers of a language learn to attend only to those features which are distinctive and to ignore those which are nondistinctive.

The subjects in the present experiment were given the difficult task of discriminating between contrasts in which the critical features were distinctive in French but not in English--the intention of the AD Test. The contrasts, however, were not as clear-cut as implied by the grouping of stimuli into three subtests. Subtest 1 presents the same contrast embedded in strings of three and four syllables; here, in all the stimuli within the subtest, the feature of rounding is superimposed on only one vowel configuration, the front high vowel. In the other subtests, however, the feature of nasality is superimposed on different vowel configurations. The articulatory characteristics of these vowel configurations

may very well operate in such a way as to modify the degree of contrast for a speaker of English. Thus, for example, the difference between /æ/ and /æ̃/ may be less noticeable than the difference between /o/ and /õ/ and, in Subtest 3, the difference between /œ̃/ and /æ̃/ may be more noticeable than the difference between /ã/ and /õ/. It has been shown that the quality of nasalization is manifested variously in the acoustic spectrum, depending on the vowel which it modifies. The coupling in of the nasal resonators results in a complex of changes rather than a simple additive one. It can be assumed that these changes are more or less detectable with different vowels (Curtis, 1968). If this be the case, then the discrimination task presented in this experiment might have been compounded to such a degree as to contribute substantially to the unreliability of the AD Test.

Furthermore, the very nature of the test focused attention on the strongly marked features of rounding and nasality. Thus, even though a subject may have been able to discriminate between contrasts (/ã/ vs. /õ/, for example), he may very well still have identified them as the same sound on the grounds that they shared the strongly marked feature to which he was attending.

One of the most persistent problems in research on auditory discrimination has been the confusion between the ability to discriminate (i.e., to detect differences between two stimuli) and the ability to identify or categorize (i.e., to label each of two stimuli as a member of a class). To understand more clearly the relationship between discrimination and identification it is helpful to refer to recent research on synthetic speech--a technique which permits the manipulation of the acoustic signal along a continuum. Work in synthetic speech by Liberman and his associates at Haskins Laboratory (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967) has shown that human perception operates differently with different classes of sounds on these two tasks.

There are speech sounds (vowels, for example) for which slight differences between two samples are easily perceived. Even so, the two different samples may both be judged to represent the "same" sound. Such sounds are commonly discriminated as "different" and identified as "same."

For other speech sounds (stop consonants, for example), slight differences are not perceived all along a continuum of physical change, as with vowels, but only at critical points on the continuum. At those points, a given slight difference (no greater than any before) will be perceived and will result in a different identification or label. These two distinct types of speech stimuli have been differentiated as "continuous stimuli" (vowels) and "categorical stimuli" (certain consonants).

> Categorical perception refers to a mode by which stimuli are responded to, and can only be responded to in absolute terms. Successive stimuli drawn from a physical continuum are not perceived as forming a continuum, but as members of discrete categories (Studdert-Kennedy, Liberman, Harris, & Cooper, in press).

On the other hand, when asked to discriminate between pairs of "continuous stimuli" (vowels),

> subjects discriminate equally well across the entire continuum: that is to say, they discriminate just as well between stimuli drawn from the same category as between stimuli drawn from different categories. . . . The process of discrimination is independent of the process of identification (p. 142).

It is apparent, then, that a test identified as an auditory discrimination test may, in fact, be measuring a complex of factors. This would account to a great degree for the lack of internal consistency evident in Tables 1 through 4.

The "Not sure" option was included in this test on the assumption that it would serve to identify instances of guessing and thus isolate much of the error variance. It was assumed that if a student had not heard one or more of the members of a triad because of some external distraction or because his attention had wandered momentarily, he would have guessed at the answer. It was also assumed that if the student had not been able to discriminate between the more difficult contrasts, he would have guessed. The "Not sure" option was expected to replace these guesses and, at the same time, to reduce the tension of the students taking the test. The latter expectation, however, was not realized. The evidence presented by the item analyses (Tables 1 through 4) and by a careful survey of the response patterns in respect to the "Not sure" option indicates that the inclusion of this option served either to encourage indecision on

the part of the students or to offer an acceptable means of avoiding a difficult decision. Thus, some of the most important differences in the abilities of the students to make correct discriminations between the vowel phonemes may have been obscured by the "Not sure" option. Much more investigation into the research concerning listener bias and response uncertainty is needed before the use made of such an option can be interpreted. In the meantime, it may be more important in a study of this kind to force a dichotomous decision for the sake of increased sensitivity on the more difficult contrast measures.

Echo Response and Recall Tests

Results of the main dependent variable (Echo Response #6) and of the Recall Test (the final test instrument in the battery) were obtained by means of the same computer program used for the AD Test. The reliability coefficient (Cronbach's Alpha) for Echo Response #6 was .43 while the reliability coefficient for the Recall Test was .51. Statistics on the individual items on ER #6 and on the Recall Test are summarized in Tables 2 and 3 respectively.

TABLE 2

Summary Statistics of Echo Response #6

Item	r_{pb}	p	sd	Skewness	Kurtosis
1. le sou	.33	.60	.49	-0.43	-1.82
2. le bateau	.30	.68	.47	-0.75	-1.44
3. le lit	.38	.82	.38	-1.71	.92
4. le mur	.32	.36	.48	.58	-1.66
5. la feuille	.36	.40	.49	.39	-1.85
6. l'agent	.50	.61	.49	-0.45	-1.80
7. la clef	.37	.64	.48	-0.58	-1.66
8. le feu	.38	.66	.47	-0.70	-1.51
9. le seau	.41	.61	.49	-0.45	-1.80
10. le train	.30	.54	.50	-0.14	-1.98
11. le pont	.54	.64	.48	-0.58	-1.66
12. le coeur	.24	.83	.38	-1.76	1.08

TABLE 3

Summary Statistics of Recall Test

Item	r_{pb}	p	sd	Skewness	Kurtosis
1. le sou	.40	.17	.38	1.76	1.09
2. le bateau	.44	.36	.48	.58	-1.66
3. le lit	.29	.22	.42	1.32	-0.26
4. le mur	.34	.52	.50	-0.08	-1.99
5. la feuille	.32	.18	.39	1.62	.63
6. l'agent	.40	.30	.46	.90	-1.19
7. la clef	.44	.56	.50	-0.24	-1.94
8. le feu	.47	.50	.50	.00	-2.00
9. le seau	.30	.20	.40	1.50	.25
10. le train	.50	.47	.50	.12	-1.98
11. le pont	.36	.30	.46	.87	-1.24
12. le coeur	.46	.58	.50	-0.30	-1.91

One of the factors that may have contributed to the low reliability coefficients of these tests is the presence of the French uvular /r/ in le mur, le train, and le coeur. It is quite possible that the difficulty of producing this sound may have influenced the students' pronunciation of the critical vowel sounds or that the American English /r/ coarticulation effects may have affected the rater's judgment of the acceptability of the adjacent vowel. Since the presence of the uvular /r/ is not a variable of concern in this experiment, these three vocabulary items could be replaced.

Another problem, and a major one in both the Echo Response and Recall Tests, is the lack of sufficient items containing the vowel sounds of interest. In the 12-item test, only four items contained a front-rounded vowel, and only three items contained a nasal vowel. The remaining five items, containing monophthongal vowels which speakers of English tend to pronounce as diphthongs, were not, in fact, relevant to the problem and

tended to obscure the issue. These items could be replaced by items containing rounded and nasal vowels and the test could then be constructed to contain multiple examples of each.

Data Analysis

To test the hypothesis that at different levels of auditory discrimination, two treatments in pronunciation training are differentially effective, scores on the main dependent variable, Echo Response #6, of all students in each treatment group were regressed on their scores in the Auditory Discrimination Test.

In view of the availability of the "Not sure" option, three methods of scoring the AD Test were considered: An uncorrected score would consider the "Not sure" response as incorrect and would assign a one to a correct response and a zero to an incorrect or "Not sure" response; a corrected score would assign a value of .50 to a "Not sure" response; and a proportion score would be derived from the number of correct responses divided by the number of items in the test (29) minus the "Not sure" responses. Consideration of these three scoring methods indicates that, for the purposes of measuring the relative performance of students, it makes little difference which method of scoring is used. This is shown in Table 4 in which the statistics relevant to the regression analyses on all three methods of scoring are presented. In all three cases, the nonsignificance of the F-ratio in the test for parallelism of regression lines indicates that the lines are parallel and that there is no interaction. Thus, the data obtained from the AD Test and the ER Test do not support the hypothesis. This is illustrated in Figure 1, in which the two regression lines and a scatter plot of the distribution of scores are presented. (Since the results from the three methods of scoring are so similar, the regressions on the uncorrected scores will be the only ones presented in this report.)

It is likely that a regression analysis based only on treatment groups might tend to obscure the differences between the performances of male and female students. Similar analyses were therefore done on the basis of sex within treatment groups. The regression lines and scatter

TABLE 4

Summary Statistics of Analyses of Regression of ER #6
on Three Types of AD Test Scores:
Both Treatment Groups

Type of Scoring on AD Test	Treatment Groups	AD Test Mean	AD Test sd	ER Test Mean	ER Test sd	r	s_e	b	F-ratio df = 1,196
Uncorrected (N - wrong)	Tape Rec. (N = 100)	17.1	3.8	7.3	2.2	.22	2.19	.13	2.00 (NS)
	Teacher (N = 100)	16.9	3.8	7.6	2.0	.04	1.99	.02	
Corrected (Right + .50 n/s)	Tape Rec. (N = 100)	19.7	3.0	7.3	2.2	.23	2.18	.17	1.16 (NS)
	Teacher (N = 100)	19.8	3.0	7.6	2.0	.10	1.98	.07	
Proportion (Right ÷ [N - n/s])	Tape Rec. (N = 100)	71.2	12.0	7.3	2.2	.20	2.20	.04	.58 (NS)
	Teacher (N = 100)	72.7	12.1	7.6	2.2	.12	1.98	.02	

plots obtained from these analyses are presented in Figures 2 and 3. The statistics relevant to these analyses are presented in Table 5. It is quite clear from the differences in the slopes of the regression lines in both treatment groups and from the significant F-ratios that there is some interaction between uncorrected Auditory Discrimination Test scores and sex. On the basis of this analysis, it would be possible to generate the following working hypothesis: At different levels of auditory discrimination, different types of pronunciation training are differentially effective for boys and girls.

The information obtained from a regression analysis is, however, somewhat limited in scope. To determine more precisely the sources of variation on the AD Test and the ER Test, the data were submitted to a three-way analysis of variance on the dimensions of sex, treatment, and *level of auditory discrimination.* For this purpose, students were assigned to high AD and low AD groups on the basis of their uncorrected AD Test scores, with the interval between 17 and 18 as the cut-off point. The results of the analysis of variance are presented in Table 6. From these results, it is evident that, although there is no significant main effect, there is a significant interaction of auditory discrimination level and sex. No effects of interaction of treatment with either of these variables is discernible. To interpret the interaction of sex and auditory discrimination level, mean scores for all groups on ER #6 are printed in Table 7.

Although the differences are small and can only suggest a trend, it is important to note that the high AD boys perform better than the low AD boys, while the opposite is true for the girls. Thus, although it might be expected that the girls with a high level of auditory discrimination would achieve the highest scores on the Echo Response Test, the means indicate an unexpected departure from predicted behavior.

It has long been a commonplace assumption among foreign language teachers that girls are better language students than boys. The two most frequently expressed reasons are that girls mature faster and, up to a certain age, learn more quickly than boys and that girls have a more positive attitude than boys toward language learning. This latter reason,

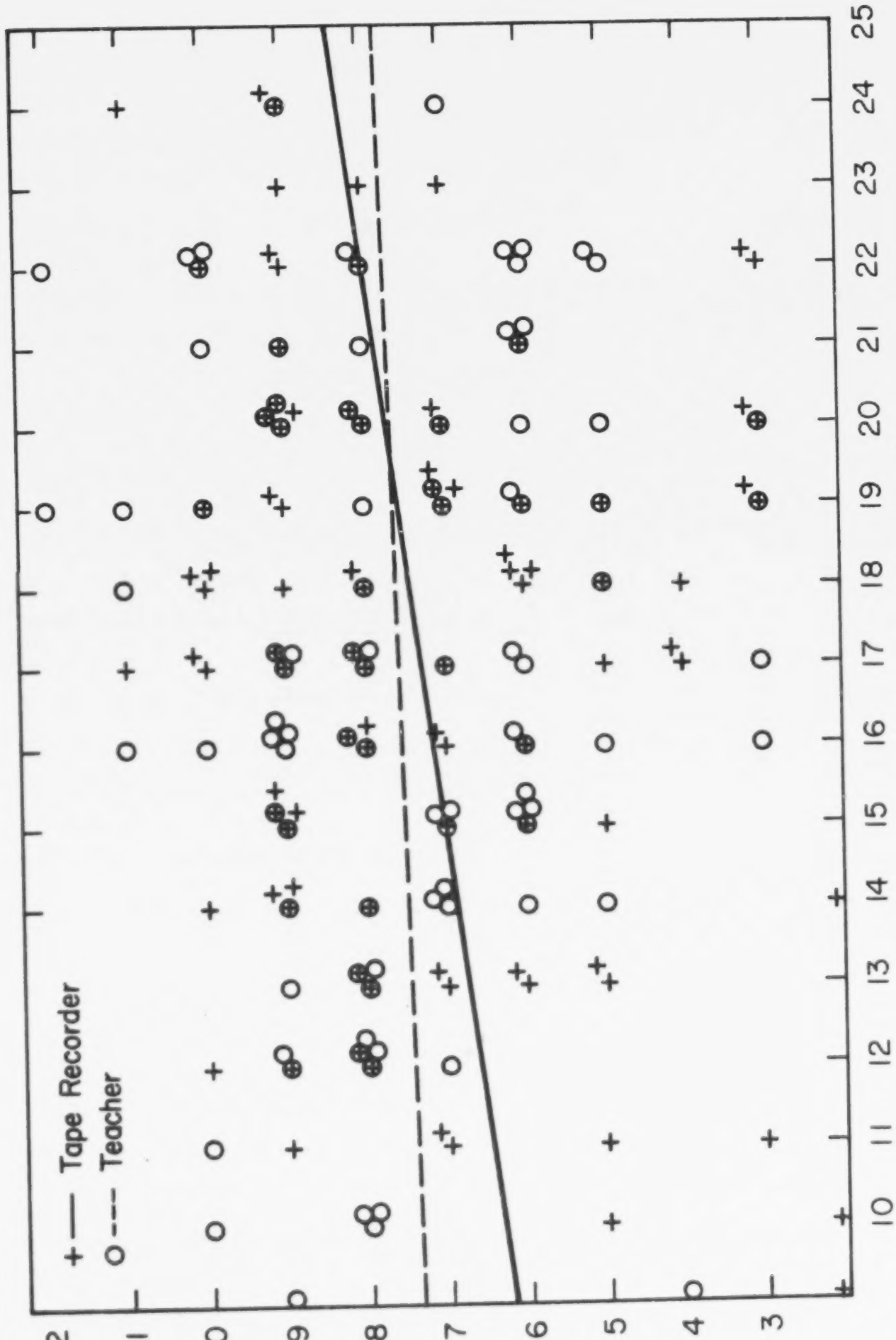

Fig. 1. Regression and scatter plot of Echo Response Test scores on AD Test (uncorrected) scores: Tape Recorder Treatment and Teacher Treatment.

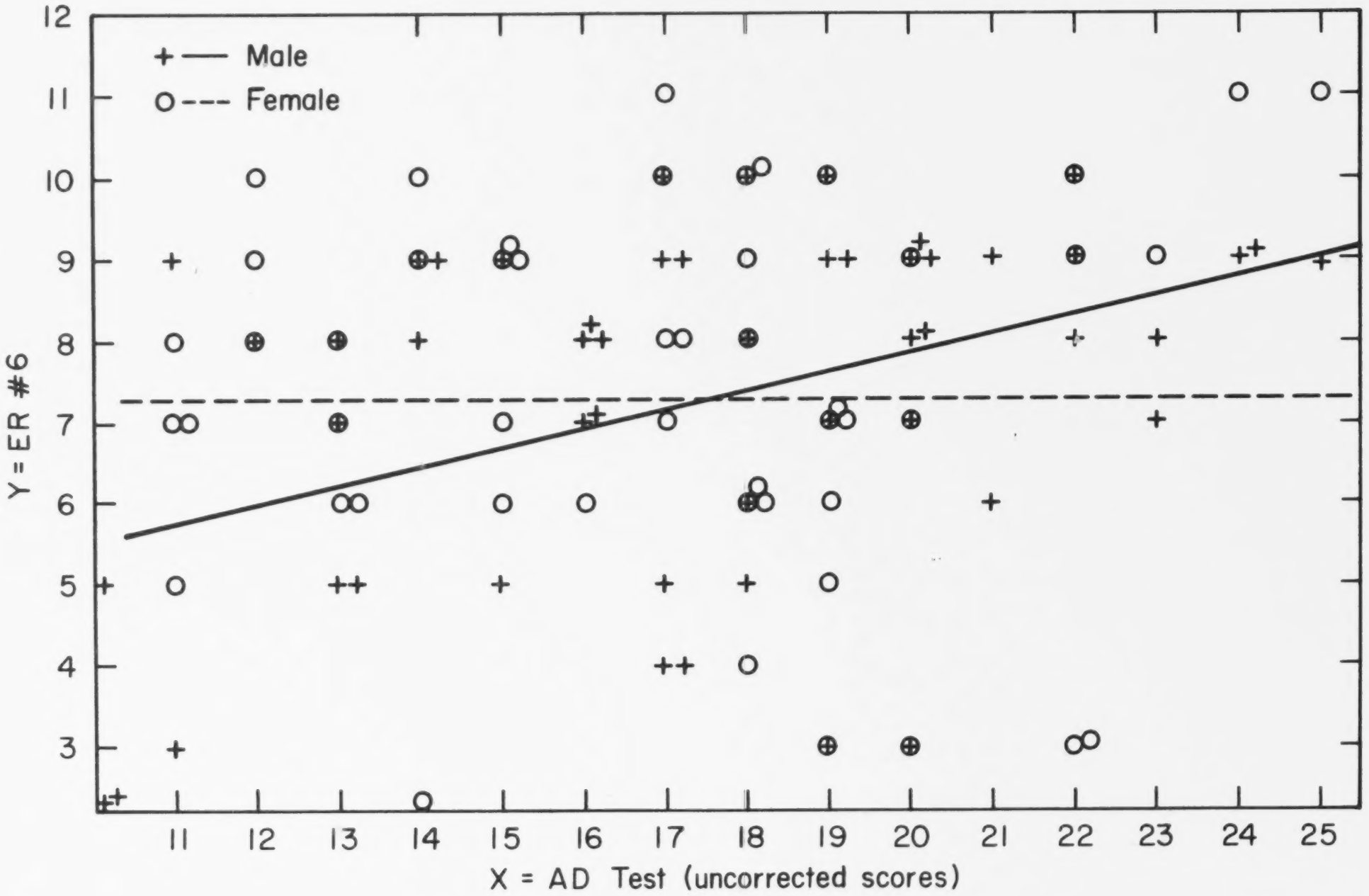

Fig. 2. Regression and scatter plot of Echo Response Test scores on AD Test (uncorrected) scores: Tape Recorder Treatment--male and female.

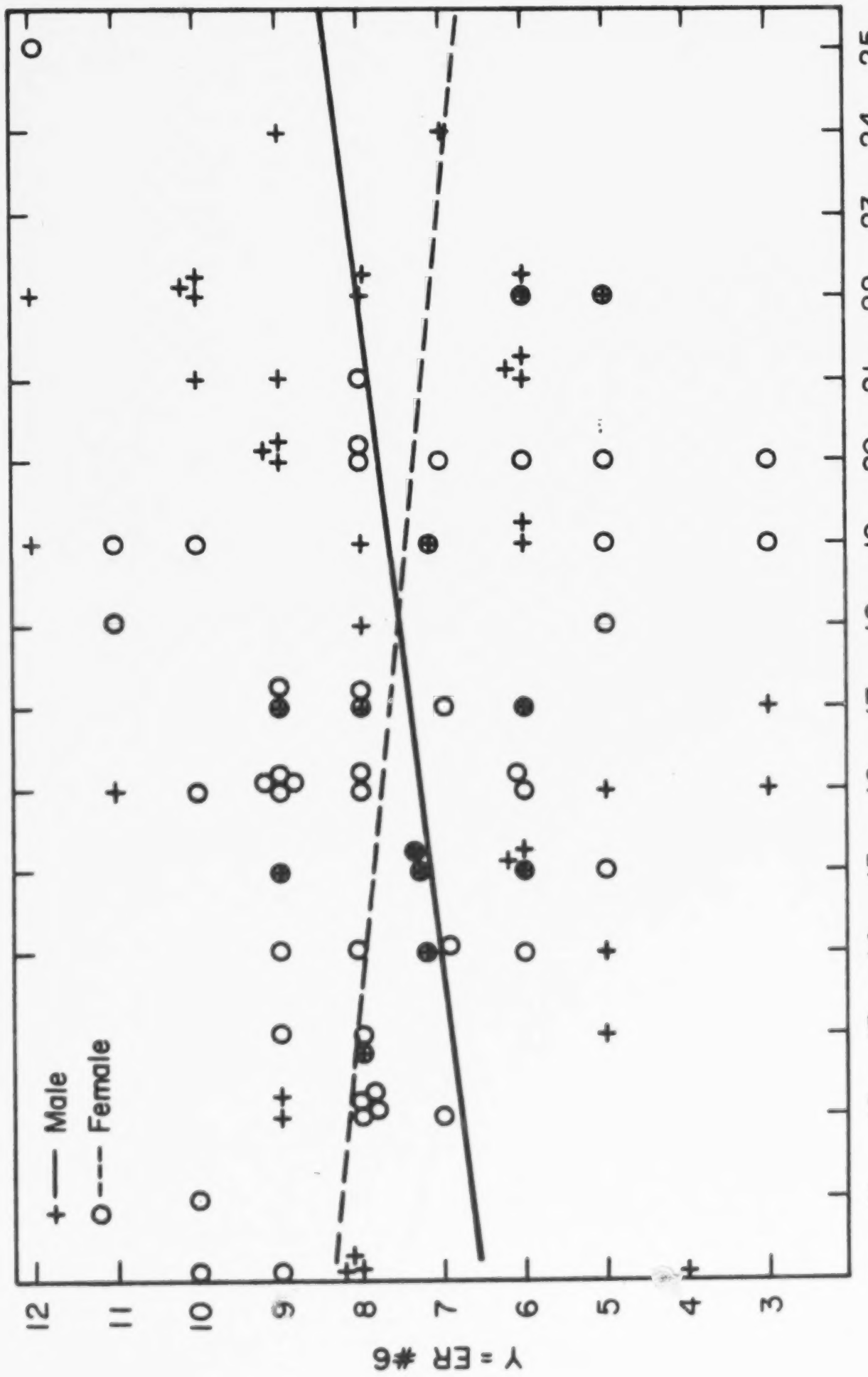

Fig. 3. Regression and scatter plot of Echo Response Test scores on AD Test (uncorrected scores: Teacher Treatment--male and female.

TABLE 5

Summary Statistics of Analysis of Regression of ER #6
on Uncorrected AD Test Scores:
Sex Within Treatment Groups

Treatment Groups	Sex	AD Test Mean	AD Test sd	ER Test Mean	ER Test sd	r	s_e	b	F-ratio df = 1.96
Tape Recorder	Male (N = 52)	17.4	3.9	7.2	2.2	.42	2.02	.24	4.29*
	Female (N = 48)	16.8	3.6	7.3	2.3	.00	2.35	.00	
Teacher	Male (N = 48)	17.6	4.1	7.5	2.1	.24	2.09	.12	4.38*
	Female (N = 52)	16.3	3.4	7.6	2.1	-.18	1.89	-.10	

* $p < .05$

TABLE 6

Three-way Analysis of Variance of Scores on Echo Response #6

Source of Variation	df	SS	ms	F-ratio
Auditory Discrimination	1	4.12	4.12	.93
Sex	1	2.25	2.25	.51
Treatment	1	3.49	3.49	.79
Auditory Discrimination x Sex	1	24.94	24.94	5.63*
Auditory Discrimination x Treatment	1	1.24	1.24	.28
Sex x Treatment	1	.01	.01	.00
Auditory Discrimination x Treatment	1	.59	.59	.15
Error	192	850.72	4.43	

* p > .05

although unsubstantiated, may indeed be valid for French, which has traditionally been considered a language for girls to study. The former reason seems to be reinforced by a recent study conducted by Asher and Garcia (1969) in which 71 Cuban immigrant children ranging in age from 7 to 19 were tested for acquisition of near-native pronunciation in English. The findings indicated that girls outranked boys two-to-one in acquisition of near-native pronunciation and that, generally, boys needed more time to acquire the phonology of the new language than did girls.

If girls learn at a faster rate than boys and if, as it is supposed, girls are more highly motivated toward learning French than boys are, the question arises why the high AD girls did not perform as well as might have been expected on the Echo Response Test. One answer may perhaps be found in the results of the Recall Test. To investigate this possibility, scores on the Recall Test were analyzed by the same three-way analysis of variance as scores on the Echo Responses. This analysis, presented in Table 8, shows no significant interactions, but does show a significant

TABLE 7

Mean Scores on Echo Response #6 for All Groups

	Tape Recorder			Teacher		
	N	Mean	sd	N	Mean	sd
Male						
Low AD Group	20	6.60	2.3	19	6.89	2.0
High AD Group	32	7.66	2.1	29	7.86	2.1
Female						
Low AD Group	20	7.40	1.9	29	7.96	1.3
High AD Group	28	7.25	2.5	23	7.26	2.4

TABLE 8

Three-way Analysis of Variance of Scores on Recall Test

Source of Variation	df	SS	ms	F-ratio
Auditory Discrimination	1	27.48	27.48	6.06*
Sex	1	19.94	19.94	4.40*
Treatment	1	7.75	7.75	1.71
Auditory Discrimination x Sex	1	.31	.31	.07
Auditory Discrimination x Treatment	1	1.84	1.84	.41
Sex x Treatment	1	.79	.79	.18
Auditory Discrimination x Sex x Treatment	1	.79	.79	.18
Error	192	870.54	4.53	.18

* $p < .05$

main effect of sex (F-ratio = 4.40). The mean scores for all groups, summarized in Table 9, show a small but consistent superiority of all subjects in the high AD groups over those in the low AD groups and a superiority of female subjects over male subjects in each of the treatment groups as well as in the AD groups.

TABLE 9

Mean Scores on Recall Test for All Groups

	Tape Recorder			Teacher		
	N	Mean	sd	N	Mean	sd
Male						
Low AD Group	20	3.70	2.0	19	3.58	1.3
High AD Group	32	4.47	2.5	29	4.48	1.9
Female						
Low AD Group	20	4.90	2.4	29	3.83	1.9
High AD Group	28	5.25	2.4	23	4.83	2.1

The difference between the results obtained in the Echo Response Test and the Recall Test points up the crucial difference between mimicry and reproduction, that is, between the imitation of sounds and the recall of sounds. Neisser (1966), in his exhaustive review of the literature, points out that it is characteristic of active verbal memory that information is recoded rather than simply echoed and cites a number of experiments indicating that the information stored in short-term memory is auditory information. In studies of substitution errors and interference errors, for example, it was found that even if the stimuli are presented visually, the substitution errors tend to involve units that sound alike

and that items which interfere with each other consistently tend to have phonemes in common. He concludes that:

> On the one hand, active verbal memory cannot be simply echoic, since what is remembered has typically been grouped and recoded until it is quite different from what was presented. On the other hand, the material is evidently stored in a form similar to that in which words are perceived, since the same confusions occur in both cases. This seeming paradox is resolved if both speech perception and verbal memory involve the same active process of synthesis, or auditory attention (p. 225).

Pimsleur (1963), in discussing his discrimination training experiment, asserts that the imitation of sounds by making certain articulatory adjustments is a motor skill and has nothing to do with categorization.

Subjects in the present study were told that they were to learn certain vocabulary items by imitating the pronunciation of the model as they looked at the picture. The goal, therefore, was correct recall rather than correct pronunciation. If, in order to attain this goal, subjects were attending to the task of grouping and recoding the auditory stimuli, it is quite probable that they were paying less attention to the motor task of making precise articulatory adjustments. It may be that the high AD girls concentrated more on identifying than on mimicking and, thus, had more opportunity to organize the stimuli into categories for the purpose of recall. It is also possible that the high AD girls, attending to the recall task, felt that they had learned the words early in the training and simply paid only cursory attention to the pronunciation task. These are obviously conjectures and cannot be substantiated by any of the obtained data, but it is evident that there are important differences between the pronunciation task and the recall task. The conflict of the two tasks may very well have served to obscure rather than reveal any relationship between the Auditory Discrimination Test and pronunciation.

Summary and Conclusions

The main purpose of this study was to investigate the validity of the assumption that auditory discrimination and pronunciation in a foreign language are closely related. If this assumption is well-founded, then it might be possible for foreign language students with high auditory discrim-

ination ability to work alone in the language laboratory with appropriately structured materials and achieve criterion pronunciation by comparing their own speech with that of the taped model. If this is possible, a reliable test of auditory discrimination might be a valuable diagnostic and predictive instrument for assigning some students to individualized language laboratory-based instruction and others to teacher-based instruction.

The design of the study involved the development and testing of an instrument for the measurement of auditory discrimination ability. This Auditory Discrimination Test presented students with 29 minimal pairs of contrasting French vowel phonemes (rounded vs. unrounded; nasal vs. non-nasal; nasal vs. other nasal) embedded in strings of three, four, or five syllables. The test was recorded on audio tape by the experimenter and presented to students free field with a portable tape recorder in their respective classrooms. The response design was in the form of ABX triads and the task of the student was to signify on an answer sheet whether he heard X as sounding more like A or B.

A second instrument, the Echo Response Test, was used to investigate the differential effects of two types of pronunciation training upon students with varying degrees of auditory discrimination ability as measured by the AD Test. The two types of pronunciation training, designed to parallel language laboratory-based and teacher-based instruction, involved the oral presentation of 12 French vocabulary items to 100 male and 100 female subjects at the sixth-grade level. In the first training condition, stimuli were presented directly by a trained investigator, while in the second condition, stimuli were presented by the tape-recorded voice of the same investigator. To eliminate experimenter bias, male and female subjects were divided equally between two investigators. In each condition, each item, accompanied by an appropriate illustration, was presented six times, with an echo response required of the subject after each presentation. After the fourth, eighth, and twelfth items, subjects were shown the illustrations in the same order in which they had been originally presented and were asked to recall the French name of the item.

To test the hypothesis that at different levels of auditory discrimination two treatments in pronunciation are differentially effective, scores on the main dependent variable, Echo Response #6, of all subjects in each treatment group were regressed on their scores on the Auditory Discrimination Test. This regression analysis revealed no differential effects of the two pronunciation training treatments constructed for this study. The hypothesis was, therefore, not supported by the data. This finding may not be generalizable to the classroom and the language laboratory, however, since there is the possibility that the intended parallel with language laboratory-based and teacher-based instruction was not realized. The constraints placed upon feedback from the experimenters in the Teacher Treatment may have eliminated the most important difference between the two conditions.

When the treatment groups were divided into male and female subgroups, the regression analyses indicated a significantly different reaction to the pronunciation training. In both training conditions, the slopes of the regression lines for the boys were substantially positive, indicating a clear relationship between auditory discrimination ability, as measured here, and pronunciation after training. However, in the Tape-Recorder Treatment, the slope of the regression line for the girls was flat, while the slope of the regression line for the girls in the Teacher Treatment was substantially negative, indicating no clear relationship of auditory discrimination ability and pronunciation after training.

These differential effects were inspected more closely by means of a three-way analysis of variance, which indicated that the interaction of sex and auditory discrimination ability contributed significantly to the variance of the Echo Response scores. In both training conditions, boys with high AD Test scores performed better on the Echo Response than the boys with low AD Test scores. On the other hand, girls with low AD Test scores performed as well as girls with high AD Test scores in the Tape Recorder Treatment and better than the girls with high AD Test scores in the Teacher Treatment.

In contrast, a similar analysis of the Recall Test showed no interaction effects, but did show a significant main effect of auditory

discrimination level and a significant main effect of sex of subjects. The mean scores for all groups showed a small but consistent superiority of all subjects in the high AD groups to those in the low AD groups and a superiority of female subjects over male subjects in each of the treatment groups as well as in the AD groups. Thus, although the performance of the boys in both tasks, pronunciation and recall, followed the predicted pattern, that of the girls in the pronunciation task did not. It is quite clear from these results that the interpretations of findings in an experiment such as this must be subject to careful considerations of the complex interactions of the sex and abilities of subjects with the nature of the experimental task.

The results of this study also indicate two theoretical distinctions which are of major importance to research on language learning and the effectiveness of the language laboratory. Comparison of the results on the Echo Response Test, which is essentially a test of mimicry, with those on the Recall Test, which is essentially a test of reproduction of meaningful sounds, reflects the crucial difference between mimicry and reproduction in foreign language learning. The reproduction of sounds associated with meaning (and separated in time from a model) gives evidence of the recoding and storing of auditory stimuli and of their availability for future use. Mimicry (temporally bound to a model) gives no such assurance.

Similarly, comparison of the several perceptual tasks involved in responding to the stimuli presented in this study underscores the necessity for distinguishing between the processes of discrimination and identification. One of the most persistent problems in research on auditory discrimination has been the confusion between the ability to discriminate (that is, to detect differences between two stimuli) and the ability to categorize (that is, to identify or to label each of two stimuli as members of a class). Recent research has revealed that the perceptual functions of discrimination and identification are not identical but, in fact, operate in a different fashion for vowels as opposed to consonants.

It is clear that both the discrimination and identification functions are affected by the process of language learning. In the learning of the

native language, a variety of contextual and distributional cues have operated to determine whether two vowel sounds which differ phonetically do or do not belong to the same phonemic category. Thus, a speaker of English can perceive a phonetic difference between a monophthongal [o] and a diphthongal [o^{w}], but will "ignore" this difference and correctly label both sounds as members of the same category. In the learning of a foreign language, however, before such cues become internalized, the learner does not yet know whether to attend to these perceptible differences or to ignore them. In such a case, members of different categories, differentiated by one or more phonetic features, may or may not be identified as different. For example, when members of phonetically similar but different categories share an overriding marked feature (such as nasality or rounding) they may be identified and labeled as being the "same."

Tests of auditory discrimination and of sound identification serve different purposes. A given test identified as an Auditory Discrimination Test or a Sound Identification Test may, in fact, be measuring a complex of factors. The predictive and diagnostic value of either test can be investigated only after appropriate instruments measuring these separate functions have been devised.

The relationships between mimicry and reproduction, between phonetic and phonemic differentiation and between the discrimination and identification of sounds are but a few of the many problems that have complicated this study and virtually all research in the field of auditory discrimination. These interrelationships have important implications with respect to research on the assumptions underlying the effective use of the language laboratory. If research in this field is to be productive and meaningful, it must be based on a better understanding of these distinctions.

References

Asher, James J., and Ramiro Garcia. "The Optimal Age to Learn a Foreign Language," The Modern Language Journal, LIII, No. 5 (May 1969), 334-341.

Brooks, Nelson. Language and Language Learning: Theory and Practice. 2d ed. New York: Harcourt, Brace and World, 1964.

Campbell, Donald T., and Julian C. Stanley. "Experimental and Quasi-Experimental Designs for Research in Teaching," in Handbook of Research on Teaching, ed. N. L. Gage. Chicago: Rand, McNally and Co., 1963.

Catford, J. C. A Linguistic Theory of Translation. London: Oxford University Press, 1965.

Curtis, James F. "Acoustics of Speech Production and Nasalization," in Cleft Palate and Communication, ed. D. C. Spriestersbach and Dorothy Sherman. New York: Academic Press, 1968.

Hebb, D. O. Organization of Behavior. New York: John Wiley and Sons, 1949.

Hutchinson, Joseph C. "The Language Laboratory: Equipment and Utilization," in Trends in Language Teaching, ed. Albert Valdman. New York: McGraw-Hill, Inc., 1966.

Jakobson, Roman, Gunnar Fant, and Morris Halle. Preliminaries to Speech Analysis. Technical Report 13. Acoustics Laboratory, Massachusetts Institute of Technology, 1952.

Keating, Raymond. "A Study of the Effectiveness of Language Laboratories." New York Institute of Administrative Research. New York: Columbia University Teachers College, 1963.

Lado, Robert. Language Testing. London: Longmans, Green and Co., Ltd., 1961.

Lane, Harlan. "Identification, Discrimination, Translation; the Effects of Mapping Physical Continua into Phoneme and Semene Categories," International Review of Applied Linguistics in Language Teaching, IV (1966), 215-226.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy. "Perception of the Speech Code," Psychological Review, LXXIV, No. 6 (November 1967), 431-461.

Neisser, Ulrich. Cognitive Psychology. New York: Appleton-Century Crofts, 1966.

Pimsleur, Paul. "Discrimination Training in the Teaching of French Pronunciation," The Modern Language Journal, XLVII, No. 5 (May 1963), 199-203.

Pimsleur, Paul. Language Aptitude Battery. New York: Harcourt, Brace, and World, Inc., 1966.

Pimsleur, P., L. Mace, and E. Keislar. Preliminary Discrimination Training in the Teaching of French Pronunciation. Final Report, United States Office of Education, Contract DHEW: SAE 8950. Los Angeles: University of California, 1961.

Politzer, R. L., and Louis Weiss. "Developmental Aspects of Auditory Discrimination, Echo Response and Recall," The Modern Language Journal, LIII, No. 2 (February 1969), 75-85.

Rivers, Wilga M. The Psychologist and the Foreign Language Teacher. Chicago: University of Chicago Press, 1964.

Scherer, George A. C., and Michael Wertheimer. A Psycholinguistic Experiment in Foreign Language Teaching. New York: McGraw-Hill Book Co., 1964.

Smith, Philip D., Jr., and Helmut A. Baranyi. A Comparison Study of the Effectiveness of the Traditional and Audio-lingual Approaches to Foreign Language Instruction Utilizing Laboratory Equipment. Final Report, Project No. 7-1033, United States Office of Health, Education and Welfare (October 1968).

Smith, Philip D., and Emanuel Berger. An Assessment of Three Foreign Language Teaching Strategies Utilizing Three Language Laboratory Systems. Final Report, Project 5-0683, United States Office of Health, Education and Welfare (January 1968).

Studdert-Kennedy, Michael, A. M. Liberman, K. S. Harris, and F. S. Cooper. "The Motor Theory of Speech Perception, a Reply to Lane's Critical Review," Psychological Review (in press).

Zeldner, Max. "The Bewildered Modern Language Teacher," The Modern Language Journal, XLVII, No. 6 (October 1963), 245-253.

APPENDIX A

Analyses of Variance of Scores on Lorge-Thorndike IQ Test, Auditory Discrimination Test and Echo Response #1

Source of Variation	df	SS	ms	F-ratio
Lorge-Thorndike IQ Test				
Class	10	3336.00	333.60	1.52
Treatment	1	219.74	219.74	.99
Class x Treatment	10	1773.00	177.30	.81
Error	178	36921.88	219.77	
Auditory Discrimination Test				
Class	10	221.01	22.10	1.59
Treatment	1	.55	.55	.04
Class x Treatment	10	164.98	16.49	1.19
Error	178	2468.71	13.87	
Echo Response #1				
Class	10	62.78	6.28	1.47
Treatment	1	.82	.82	.19
Class x Treatment	10	58.20	5.82	1.36
Error	178	761.88	4.28	

APPENDIX B

Auditory Discrimination Test

Tape Script

Practice words: tãte - tɔte - tãte
sɛro - søro - søro

Here are the first three items. Listen carefully and mark your choice:

1. lãpa - lõpa - lõpa

2. pãse - pose - pãse

3. tõbe - tobe - tõbe

The next set will be a little harder, because the words are longer. Listen very carefully. Listen.

4. desæpa - desæ̃pa - desæpa

5. pasifi - pasyfi - pasifi

6. ralãti - ralõti - ralõti

Now, listen carefully to the next set of words and mark your choice. Listen.

7. dytãta - dytɔta - dytɔta

8. labõte - labote - labote

9. desiʒe - desyʒe - desiʒe

10. dylœ̃pa - dylæ̃pa - dylæ̃pa

All right. The next items will have some words a little longer than the previous ones. Listen carefully and see how well you can do. Listen.

11. desæ̃pati - desæpati - desæ̃pati

12. ləbõkare - ləbokare - ləbõkare

13. pasifiza - pasyfiza - pasyfiza

14. ralãtise - ralõtise - ralãtise

Now, put your pencils down and relax for a minute. (Pause of 30 seconds provided here.) All right, the next group of items has some long words, so listen carefully. Listen.

15. salõtɛre - salãtɛre - salõtɛre

16. dypõzofø - dypozofø - dypozofø

17. desidema - desydema - desidema

18. matãtiva - matɔtiva - matãtiva

19. kəpõsevu - kəpãsevu - kəpõsevu

Now, here is the next group of words. Listen carefully and make your choice. Listen.

20. dytãta - dytɔta-- dytɔta

21. desæ̃pati - desæpati - desæ̃pati

22. desidema - desydema - desidema

23. ralãti - ralõti - ralõti

24. ləbõkare - ləbokare - ləbõkare

Finally, here is the last group. Listen

25. desiʒe - desyʒe - desiʒe

26. matãtiva - matɔtiva - matãtiva

27. dylœ̃pa - dylæ̃pa - dylæ̃pa

28. kəpõsevu - kəpãsevu - kəpõsevu

29. pasifiza - pasyfiza - pasyfiza

APPENDIX C

Answer Sheet for Auditory Discrimination Test

Sound Matching

NAME_____

TEACHER_____

	1st	2nd	Not sure		1st	2nd	Not sure
1.	0	0	0	1.	0	0	0
2.	0	0	0	2.	0	0	0
3.	0	0	0	3.	0	0	0
				4.	0	0	0
1.	0	0	0	5.	0	0	0
2.	0	0	0				
3.	0	0	0	1.	0	0	0
				2.	0	0	0
1.	0	0	0	3.	0	0	0
2.	0	0	0	4.	0	0	0
3.	0	0	0	5.	0	0	0
4.	0	0	0				
				1.	0	0	0
1.	0	0	0	2.	0	0	0
2.	0	0	0	3.	0	0	0
3.	0	0	0	4.	0	0	0
4.	0	0	0	5.	0	0	0